Original Papers

Polish Psychological Bulletin
2016, vol. 47(1) 73–80
DOI - 10.1515/ppb-2016-0008

Bartłomiej Kroczek[*]
Michał Ociepka[*]
Adam Chuderski[**]

No Spearman's Law of Diminishing Returns for the working memory and intelligence relationship

***Abstract**: Spearman's Law of Diminishing Returns (SLODR) holds that correlation between general (g)/fluid (Gf) intelligence factor and other cognitive abilities weakens with increasing ability level. Thus, cognitive processing in low ability people is most strongly saturated by g/Gf, whereas processing in high ability people depends less on g/Gf. Numerous studies demonstrated that low g is more strongly correlated with crystallized intelligence/creativity/processing speed than is high g, however no study tested an analogous effect in the case of working memory (WM). Our aim was to investigate SLODR for the relationship between Gf and WM capacity, using a large data set from our own previous studies. We tested alternative regression models separately for three types of WM tasks that tapped short-term memory storage, attention control, and relational integration, respectively. No significant SLODR effect was found for any of these tasks. Each task shared with Gf virtually the same amount of variance in the case of low- and high-ability people. This result suggests that Gf and WM rely on one and the same (neuro)cognitive mechanism.*

***Key words**: working memory, fluid intelligence, Law of Diminishing Returns*

Fluid intelligence (reasoning ability; Gf) is the core component ability of the general factor (*g*) of human intelligence (see McGrew, 2009). Gf consists of using reasoning (inductive, deductive, spatial, etc.) to solve novel abstract problems that cannot be solved solely on the basis of one's knowledge. Gf is most often assessed with matrix problems or visual analogies (Snow, Kyllonen, & Marshalek, 1984). Gf has a profound influence on human behavior (see Deary, 2012), predicting socio-economic status of people as well as their daily life successes. Gf also strongly correlates with many cognitive abilities, like reading comprehension, creativity, learning, school achievement etc.

One influential hypothesis regarding *g*'s (and, thus, Gf's) predictive power with regard to other cognitive abilities is Spearman's (1927) *Law of Diminishing Returns* (SLODR), rediscovered in current psychology by Detterman and Daniel (1989). SLODR assumes that correlation between *g*/Gf and other cognitive abilities weakens with increasing ability level. Thus, cognitive processing in low ability people is most strongly saturated by *g*/Gf – their abilities are quite homogeneous, whereas processing in high ability people depends less on *g*/Gf – their abilities are more heterogeneous. Numerous studies confirmed the SLODR effect (e.g., Der & Deary, 2003; Evans, 1999; Jensen, 2003; Reynolds & Keith, 2007; Reynolds, Keith, & Beretvas, 2010; but see Hartmann & Reuter, 2006). For example, low *g* is more strongly correlated with crystallized intelligence than is high *g* (Reynolds & Keith, 2007), and an analogous relationship was observed for creativity (the threshold hypothesis; see Jauk, Benedek, Dunst, & Neubauer, 2013) and processing speed (Der & Deary, 2003).

Probably the strongest known neurocognitive predictor of *g*/Gf is *working memory* (WM) – the mind/brain's mechanism responsible for the active maintenance and transformation of the limited amount of information for the purpose of the current task. WM is usually tested with simple tasks that require memorization and later recognition or recall of some items (Cowan, 2001). Specifically,

[*] **Institute of Computer Science and Computational Mathematics, Jagiellonian University in Krakow, Łojasiewicza 6, 30-348 Krakow, Poland, bartek.kroczek@gmail.com, ociepka.m@gmail.com**
[**] **Institute of Philosophy, Jagiellonian University in Krakow, Grodzka 52, 31-044 Krakow, Poland, adam.chuderski@gmail.com**

Corresponding author: Adam Chuderski, Institute of Philosophy, Jagiellonian University, Grodzka 52, 31-044 Krakow, Poland, adam.chuderski@gmail.com

the average number of detected or recalled elements or relations is taken as the WM capacity (WMC). It has been demonstrated that WMC usually explains between half (Kane, Hambrick, & Conway, 2005) and three quarters (Oberauer, Schulze, Wilhelm, & Süβ, 2005) of Gf variance. Some studies have even reported that WMC is isomorphic to Gf (Chuderski, 2013; Oberauer, Süß, Wilhelm, & Wittmann, 2008; Martínez et al., 2011). However, WM is itself quite a complex construct, and much research has been devoted to understanding what mechanisms underlie its capacity.

One influential theory assumes that individual performance in both WM tasks and *g*/Gf tests depends on *attention control* exerted over cognitive processes, which includes the goal-driven directing of attention and the filtering out distraction (e.g., Burgess, Gray, Conway, & Braver, 2011; Kane & Engle, 2002). Evidence for this theory comes from demonstrations of associations between *g*/Gf and tests of attention control, such as tasks which require prolonged vigilance, coping with interference, or making fast antisaccades (e.g., Burgess et al., 2011; Unsworth, Redick, Lakey, & Young, 2010). The attention control theory of fluid intelligence holds that people with low attention control are poor test-takers because they find it difficult to maintain reasoning goals in WM, and thus their cognitive processing is prone to frequent capture by irrelevant stimuli.

Alternatively, it was shown that performance on simple *short-term memory* (STM) tasks, which require little attention control, was at least as good a predictor of *g*/Gf as performance on tasks requiring executive control, when rehearsal and chunking were blocked in the former tasks (Colom, Abad, Quiroga, Shih, & Flores-Mendoza, 2008; Cowan, Fristoe, Elliott, Brunner, & Saults, 2006; Martínez et al., 2011). Moreover, Chuderski, Taraday, Necka, and Smolen (2012) showed that tasks tapping storage could explain 70% of variance in Gf, and when differences in storage capacity were controlled, attentional control was no longer a significant predictor of Gf (for a similar result see Martinez et al., 2011). These results suggest that sheer storage capacity (the number of items simultaneously held in WM) may be the key determinant of intelligence, probably as it allows an individual to keep the subproducts of reasoning in the most active and accessible part of WM (see Carpenter, Just, & Shell, 1990; Cowan et al., 2006).

WM may also play an important role in *g*/Gf because it affects what relations can be constructed among WM items (e.g., Halford, Baker, McCredden, & Bain, 2005; Hummel & Holyoak, 2003; Viskontas, Holyoak, & Knowlton, 2005). Notably, Oberauer and his collaborators (e.g., Oberauer, Süß, Wilhelm, & Sander, 2007; Oberauer, Süβ, Wilhelm, & Wittmann, 2008) proposed that *relational integration* was crucial to intelligence. Relational integration consists of the construction of flexible, temporary bindings between a number of chunks held in WM, or between them and their corresponding mental coordinates, in order to develop a more complex, relational structure. This structure may include either concrete coordinates (e.g., serial positions in a recall task), or abstract placeholders (e.g., roles in a schema in a reasoning test). The temporary bindings allow an individual to integrate information into completely new relational structures. Oberauer et al. (2008) measured ability to process bindings using tasks which required participants to constantly monitor perceptually available stimuli so as to detect simple relations such as three rhyming words appearing in a row or column of a three-by-three matrix of words. These authors have demonstrated that scores in such tasks can be excellent predictors of WMC and *g*/Gf (see also Chuderski, 2014).

Important knowledge on why WMC so strongly predicts *g*/Gf may be acquired by means of testing whether WMC, like other cognitive abilities, is subject to SLODR. However, to date this issue has not been examined satisfactorily. Our aim consists of a preliminary investigation of the SLODR effect for the Gf-WMC relationship, using a large data set from our own previous studies. We conduct our tests separately for the three types of WM tasks, each tapping one of three crucial WM mechanisms (short-term memory storage, attention control, and relational integration) proposed by the three above mentioned accounts of WM. This seems a rational strategy, especially as some existing theoretical models predict different SLODR effects for different WM mechanisms (see below).

Four hypotheses pertaining to the relationship between WMC and *g*/Gf, depending of the *g*/Gf level, can be considered. First, the SLODR effect can be confirmed also in the case of WMC, with WMC predicting more variance in *g*/Gf for less intelligent participants than for more intelligent ones. Second (the null hypothesis), no SLODR effect can be found, meaning that WMC predicts similar part of ability variance regardless of *g*/Gf. Moreover, some evidence suggests that it is possible that the contribution of each WM mechanism to *g*/Gf depends on the strategy used for solving an intelligence test, and in consequence a given cognitive mechanism may play a more important role in *g*/Gf for some people than for others (see Conway, Getz, Macnamara, & Engel de Abreu, 2011). For example, differences in reasoning strategies on Gf tests have been observed between people scoring high versus low (Vigneau, Caissie, & Bors, 2006). A more effective solution-construction strategy that requires the simultaneous integration of all elements of the solution at the same time might rely primarily on relational integration, whereas a less effective (i.e., leading to underperformance) response-elimination strategy that is known to be prone to interference might rely more on attention control. Thus, the third hypothesis postulates that attention control predicts most *g*/Gf variance in less intelligent participants, while its contribution diminishes with increasing intelligence. The reverse would be true of relational integration and short-term memory. Finally, one seminal model of Gf (Carpenter et al., 1900) predicted that storage capacity is crucial for easier Gf test items (i.e., for the scores of low ability people), whereas the effective strategic control (e.g., management of reasoning goals, preventing distraction etc.) is necessary for the hardest items that make up scores of high ability participants. Thus, given that strategic control is somehow related to attention control, this model implies the fourth possible hypothesis, stating that short-term memory (and, possibly, relational integration) will predict the largest amount of Gf variance in low ability participants, but their contribution to Gf will decrease with increasing Gf level, whereas the reverse will be true of attentional control.

Method

Participants

A total of 610 volunteer participants (375 women and 235 men) were recruited via publicly accessible social networking websites. Each participant gave informed consent, was informed that she or he can freely leave the lab at any time, and was paid 60 Polish zloty. Six additional participants were excluded from analysis due to missing some tasks. The mean age of participants was 23.6 years (*SD* = 4.6, range 18–46).

Measures of fluid intelligence

Two paper-and-pencil tests of reasoning were used, Raven's Advanced Progressive Matrices (Raven et al., 1983), and the figural analogy test (Orzechowski & Chuderski, 2007). The 36 items of the Raven's matrices test consist of a three-by-three matrix of figural patterns in which the bottom-right pattern is missing; subjects must choose a potential match for the missing pattern from eight response options. The task is to discover the rules governing the configuration of the patterns and apply them to select the single correct response option. The analogy test consists of 36 figural analogies of the form 'A is to B as C is to X', where A, B and C are relatively simple patterns of figures. The relationship between A and B is governed by between two and five latent rules (applying to symmetry, rotation, size, color, thickness, number of objects, etc.), and X is an empty space. The task is to choose from four options the figure which is related to figure C in the same way that B is related to A. Scores on the Raven's matrices test and the analogy test were the total number of items answered correctly.

Short-term (primary) memory tasks

Two variants of an array-comparison task which is commonly believed to tap primary memory capacity (Cowan et al., 2006) were used. Both variants consisted of 90 trials. On each trial a virtual 4 × 4 array was filled with five to nine stimuli, picked from a set of ten Greek symbols (e.g., α, β, χ, and so on), or colored squares (i.e., the letter and color variants of the task, respectively). The array was presented for a period equal to the number of items multiplied by 300 ms, and then followed by a black square mask of the same size as the array, presented for 1.2 s. In a random 50% of trials, the second array was identical to the first; in the remaining trials the second array differed from the first by exactly one item in one position, which was always a new item (not a duplicate of an item from another position). When the arrays were different, the new item was highlighted with a square red border. When they were identical, a random item was highlighted in the same way. The task was to press one of two response keys to indicate whether the highlighted item was the same or different in the two arrays (maximum response latency was 4 s). The tasks were self-paced. The score was the difference between the proportion of correct responses when the arrays were different and the proportion of incorrect responses when the arrays were the same, multiplied by the set size (see Cowan et al., 2006).

Attention control tasks

Two variants of the antisaccade task were applied, which is frequently used as a measure of attention control (e.g., Unsworth, Spillers, Brewer, & McMillan, 2011). Each variant consisted of 40 self-paced trials. Each test trial consisted of four events. First, a cue was presented for 1.5 s to prompt subjects to look at the side opposite to a rapidly flashing black square. Next, a fixation point was presented in the center of the screen for 1–2 s. Then, the flashing square was shown in the middle of the left or right side of the screen, about 16 cm from the fixation point, for 0.15 s. Finally, a small dark gray arrow (pointing left, down, or right; spatial version of the task), or a string ('left', 'down', 'right'; letter version), was presented in the middle of the opposite side of the screen to the square for only 0.2 s before being replaced by a mask. The task was to look away from the flashing square in order to observe the direction of the arrow or the identity of the string and to press the associated key. The dependent variable in each task was mean accuracy.

Relation integration tasks

Participants' relational integration ability was assessed using the modified, no-memory version of the alphanumeric monitoring task, originally devised by Oberauer et al. (2008). The stimulus for each trial on the task consisted of a 3x3 array of three-symbol strings. In the letter version of the task, the strings contained three letters from a set of ten consonants; in the number version they were three-digit numbers. Depending on the task variant, participants were asked to detect whether any of the rows or columns consisted of three strings ending with the same (the three-same variant) or different digit or letter (the three-different variant; for a more detailed description of this task see Chuderski, 2014). On half the trials, the array included one of the specified configurations; on these trials participants were required to press the space key to indicate that they had detected this configuration. On the rest of the trials, the array did not contain any of the specified configurations. Trials lasted 5.5 s and were followed by a 0.1 s blink separating subsequent arrays. Successive arrays contained between one and four unchanged strings. In each version (letter or digit stimuli) of each task variant there were forty test trials. The dependent variable was the mean percentage detection of matching configurations minus the mean percentage of false alarm errors (see Snodgrass & Corwin, 1988).

Procedure

The study consisted of two experiments (346 and 264 people), separated in time by several months. Apart from the tasks above reported, depending on a subset of participants from two to twelve other tasks (e.g., working memory, attention, and concept discovery tests) were applied, but only the eight tests reported here were administered to the whole sample. Participants were tested in a cognitive psychology lab, in groups from six to twelve people. Participants completed the tests over two sessions (including either working memory tasks or intelligence tests) separated by a short break; session order was randomized across participants. As the 346-people group solved both the Raven and analogies test under no time pressure (in 60 and 45 minutes, respectively), whereas the remaining 264 people fulfilled these tests under time pressure (in 20 and 16 minutes, respectively), the standardized scores (i.e., *SD*s) were calculated for each group separately, in order to the test scores could be validly compared. Nevertheless, mean scores on speeded tests are generally close to scores on unspeeded tests (Hammel & Schmittmann, 2006).

Results

Table 1 presents descriptive statistics for measures used in the study. Table 2 shows their correlation matrix. For consecutive analyses, we standardized all the measures, and calculated four factors representing fluid intelligence (GF), short-term (primary) memory capacity (STM), control over antisaccades (CON), and relational integration (REL), using respective means of standardized measures (e.g., STM was a mean of standardized scores on the letter and color variants of the array-comparison tasks). The factors' descriptive statistics are shown in Table 1.

In general, analyzing the SLODR effects can be validly achieved by neither standard regression methods (linear regression, linear structural equation models) nor standard factorial methods (e.g., exploratory and confirmatory factor analysis). For example, the testing, by

Table 1. Descriptive Statistics for Measures Used in the Study (N = 610)

Task	*M*	*SD*	Range	Skew	Kurt.
Color arrays	3.18	1.39	-1.97–6.30	-0.54	0.12
Letter arrays	2.81	1.49	-2.05–6.40	-0.21	-0.35
Arrow antisaccade	0.68	0.25	0.02–1.00	-0.69	-0.65
Letter antisaccade	0.75	0.24	0.00–1.00	-1.13	0.27
Three–same	0.73	0.18	-0.15–1.00	-1.32	2.16
Three–different	0.37	0.24	-0.25–0.92	-0.35	-0.56
Raven (standardized)	0.00	1.00	-3.35–2.20	-0.62	0.41
Analogies (standardized)	0.00	1.00	-2.76–2.73	-0.12	-0.44
STM factor	0.01	0.99	-2.90–2.56	-0.38	-0.26
CON factor	0.00	1.00	-2.98–1.24	-0.89	-0.23
REL factor	0.01	1.00	-3.41–2.15	-0.54	-0.09
GF factor	0.00	0.90	-2.97–2.19	-0.35	-0.04

Table 2. Correlation Matrix for Measures Used in the Study

Task	1	2	3	4	5	6	7	8
1. Color arrays	**.79**							
2. Letter arrays	.52	**.83**						
3. Arrow antisaccade	.45	.44	**.93**					
4. Letter antisaccade	.44	.42	.79	**.93**				
5. Three-same	.29	.29	.24	.30	**.88**			
6. Three-different	.22	.21	.21	.26	.53	**.86**		
7. Raven	.40	.35	.36	.40	.46	.39	**.84**	
8. Analogies	.36	.31	.36	.43	.37	.35	.64	**.76**

Note. Reliabilities (Cronbach alpha) are presented in bold.

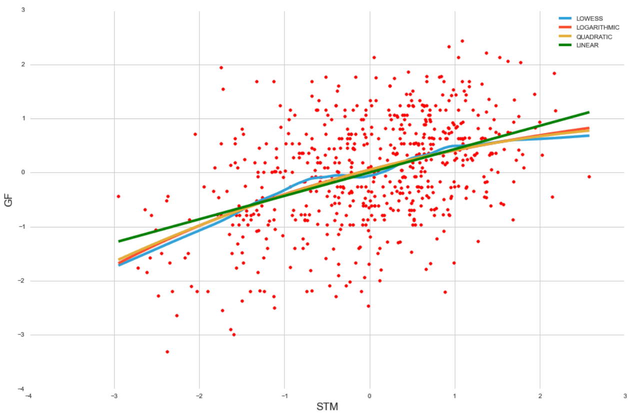
using linear regression, of how well one variable predicts another variable at different levels of the latter variable means that the latter variable appears in both the left and the right part of the regression equation at the same time. Early studies solved this problem by splitting the sample (usually at median), ideally using an intelligence subtest that was not included in the test battery investigated, and by calculating the common variance separately for two subsamples (e.g., Detterman & Daniel, 1989). This method inherits all weaknesses of the split analyses, including decreased power/reliability, increased likelihood of statistical artifacts, and arbitrariness of the split criterion (MacCallum, Zhang, Preacher, & Rucker, 2002). Newer methods use sophisticated modeling, like factor mixture modeling (Reynolds et al., 2010). Although they escape the problems of splitting, these models are quite complicated, they rely on multiple assumptions, and their clear interpretation is difficult. Thus, SLODR analyses might benefit from avoiding both splitting and surplus complexity.

Here, in order to examine possible SLODR effects, we used as simplistic analytic methods as possible, and we relied primarily on data visualization. Consequently, we applied the LOWESS (locally weighted scatterplot smoothing; see Cleveland, 1981) method in order to visualize the locally defined regression line of each WM variable on GF. LOWESS results in data smoothing based on replacing all points with the average of the neighboring data points, according to the formula:

$$y_{lowess}(i) = \frac{1}{2N+1}(y(i+N) + y(i+N-1) + \ldots + y(i-N))$$

where N is the number of neighboring data points on either side of data point $y(i)$, and $2N+1$ is the span of smoothing. In our analysis, N was arbitrarily set to 203 (⅓ of the sample). This relatively simple method allows for the continuous tracing of any changes in the relationship between regressors and regressands without the need of presetting the shape of this relationship. LOWESS can be interpreted as the saturated regression model of data that gives the best possible fit (the highest value of R^2). Then, we formally tested the regression models that assumed certain functional dependencies between a given WM factor and the GF factor, as suggested by the shape of the LOWESS line. All model parameters were fitted using the Levenberg-Marquardt algorithm. We used Akaike's (1974) Information Criterion (AIC) as the goodness-of-fit measure. AIC differences surpassing 5 indicate a significant loss of fit, its differences below 3 suggest no significant loss of fit, and differences between 3 and 5 are interpreted as inconclusive. AIC favors parsimonious models (as long as they fit well) by penalizing models for the number of their parameters, that is, by increasing its value by 2 for each additional parameter.

Figure 1. Scatter plot and the best-fitting lines of regression of the fluid intelligence factor (GF) on the short-term memory factor (STM), for the LOWESS, linear, logarithmic, and quadratic models



The LOWESS line for the STM-GF relationship is presented in Figure 1. This line suggested that the relationship in question might be nonlinear, with some attenuation of the increase in GF with increasing STM for the people showing the largest STM scores. Thus, we tested three nonlinear regression models: the logarithmic ($a+be^{-cx}$), the quadratic ($a+bx+cx^2$), and the cubic ($a+bx+cx^2+dx^3$), and contrasted them with the linear model ($a+bx$). Figure 1 shows the respective regression lines, and Table 3 includes the parameters values and fit statistics for each model. Substantially larger AIC values of the nonlinear models clearly show that their minimally better fit did not justify their greater complexity, in comparison to the – lowest in AIC value – linear model. As an additional check of the SLODR hypothesis, which originally pertained not to the shape of the best-fitting regression line, but to the amount

Table 3. Model parameters and fit statistics (Sum of Squared Residuals, R2, and Akaike Information Criterion) for regressions of Gf on STM (N = 610)

Model	*a*	*b*	*c*	*d*	SSR	R^2	AIC
LOWESS (saturated)	–	–	–	–	0.720	0.196	–
LINEAR	0.000	0.432	–	–	0.727	0.186	20.70
LOGARITHMIC	1.720	-1.670	0.241	–	0.725	0.191	29.82
QUADRATIC	0.051	0.413	-0.051	–	0.725	0.191	29.82
CUBIC	0.049	0.396	-0.046	0.007	0.725	0.191	38.60

of variance in ability explained by the regressor in the low- versus high-end of the distribution (what depends on the scattering of data around the line), we compared the R^2 statistic for the 305 below-median STM scorers ($R^2 = 0.195$) versus the 305 above-median ones ($R^2 = 0.174$). The difference between R^2 values was negligible. Thus, although the LOWESS method suggested some minor differences for the STM-GF relationship between low- and high-scoring participants, the closer inspection revealed no significant SLODR effect for this relationship.

For the CON-GF relationship, the LOWESS line yielded an almost linear shape (see Figure 2). Especially, for the logarithmic model, the *c* value was fitted to zero, thus making it a linear model. Comparison of the linear, logarithmic, quadratic, and cubic models, left the former model being the one optimally combining the sufficient fit to data with the largest parsimony (see Table 4). Again, there was no visible difference in the R^2 statistic between the 305 below-median CON scorers ($R^2 = 0.190$) versus the 305 above-median ones ($R^2 = 0.175$).

Figure 2. Scatter plot and the best-fitting lines of regression of the fluid intelligence factor (GF) on the attention control factor (CON), for the LOWESS, linear, logarithmic, and quadratic models

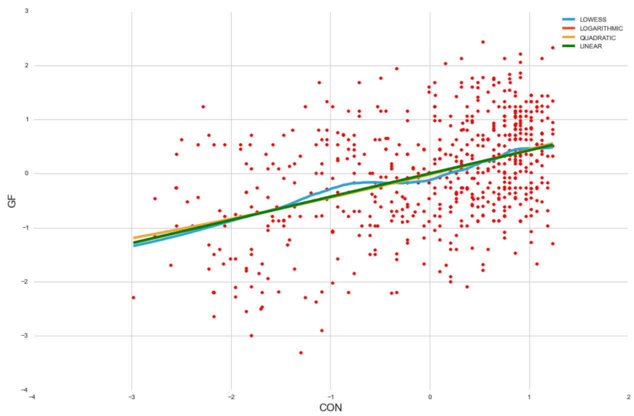


Figure 3. Scatter plot and the best-fitting lines of regression of the fluid intelligence factor (GF) on the relational integration factor (REL), for the LOWESS, linear, logarithmic, and quadratic models

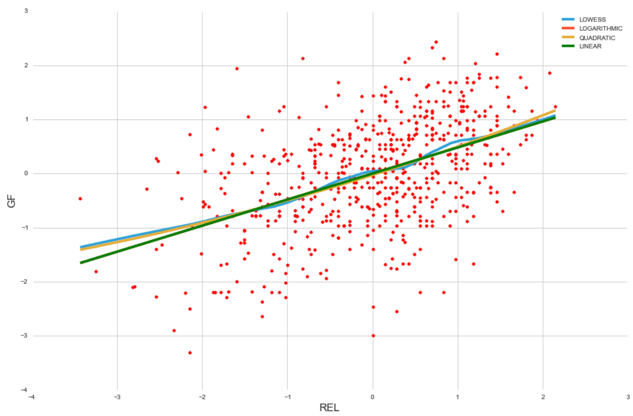


Table 4. Model parameters and fit statistics (Sum of Squared Residuals, R^2, and Akaike Information Criterion) for regressions of Gf on CON (N = 610)

Model	*a*	*b*	*c*	*d*	SSR	R^2	AIC
LOWESS (saturated)	–	–	–	–	0.728	0.190	–
LINEAR	0.000	0.429	–	–	0.732	0.184	20.71
LOGARITHMIC	2076.62	-2076.62	0.000	–	0.732	0.184	29.84
QUADRATIC	-0.017	0.444	0.017	–	0.732	0.184	29.84
CUBIC	-0.016	0.446	0.014	-0.002	0.732	0.184	38.64

Table 5. Model parameters and fit statistics (Sum of Squared Residuals, R^2, and Akaike Information Criterion) for regressions of Gf on REL (N = 610)

Model	*a*	*b*	*c*	*d*	SSR	R^2	AIC
LOWESS (saturated)	–	–	–	–	0.701	0.236	–
LINEAR	0.000	0.481	–	–	0.704	0.232	20.59
LOGARITHMIC	4190.61	-4190.61	0.000	–	0.704	0.232	29.66
QUADRATIC	-0.027	0.496	0.027	–	0.704	0.233	29.66
CUBIC	-0.026	0.501	0.025	-0.002	0.704	0.233	38.39

Basically the same results were noted for the REL-GF relationship (see Figure 3 and Table 5), with the linear model providing the best account of data. In this case, either, no detectable difference in the R^2 statistic between the 305 below-median REL scorers ($R^2 = 0.238$) versus the 305 above-median ones ($R^2 = 0.224$) could be found.

Discussion

In contrast to numerous above cited studies, which confirmed that SLODR applies to various intellectual abilities, like crystallized/verbal/learning ability, creativity, and processing speed, this study demonstrated no SLODR effect for WM as measured with the three types of tasks that tapped: short-term storage, attention control, and relational integration. No evidence for any of the three hypotheses that were alternative to the null hypothesis could be found. The pattern of data was mutually coherent, and it cannot just result from an improper measurement of Gf and WM, as the sample was substantial, our measures yielded satisfactory reliability, and the mean zero-level correlation (*r*s around .4) between Gf and WM tests resembled correlations usually observed in Gf-WM studies (see metaanalysis in Chuderski, 2013).

In our view, these results can be best understood if fluid intelligence is conceptualized simply as (i.e., it is equated to) the (overall) effectiveness of working memory. Specifically, in terms of necessary cognitive processing, fluid intelligence primarily encompasses deriving proper relations from the data given, constructing the model of a situation using these relations (as well as some counterexample models; see Goodwin & Johnson-Laird, 2005), and applying them to solve (usually novel and nontrivial) problems. For this reason, Gf is often termed reasoning ability, and is usually measured with deductive and inductive reasoning tests. Several existing models of such a kind of reasoning (e.g., Chuderski & Andrelczyk, 2015; Halford, Wilson, & Phillips, 1998; Hummel & Holyoak, 2003; Ragni & Knauff, 2013) predict that its general effectiveness (although not necessarily its specific qualitative characteristics) is determined by available WMC to a substantial extent. Thus, if both regressor and regressand represent the same ability, there is no issue of a larger or smaller loading of one factor on another (a case addressed by SLODR), as they both represent the workings of the same cognitive mechanism (so no SLODR can arise).

The above interpretation can potentially be questioned on the basis of only moderate correlations between Gf and WM measures. If we advocate for isomorphism between Gf and WM, then the observed Gf-WM correlations should approach unity. However, it must be acknowledged that single measures, or even compound measures averaging such simple measures, reflect a lot of noise and task-specific variance, and one should not expect perfect correlations in such a case. Moreover, we investigated the SLODR effect separately for each of three widely identified WM tasks (and not for general WMC representing the overall effectiveness of WM). Numerous studies demonstrated that when the strength of relationship between Gf and WMC is estimated on the level of general WM construct (a latent variable calculated from the wide range of WM tasks), instead of single WM tasks, the observed correlations often reach unity (Chuderski, 2013; Colom et al., 2008; Martinez et al., 2011; Oberauer et al., 2005; Oberauer et al., 2008). Another question pertaining to the present study might ask whether the lack of SLODR effect observed in the case of WM and Gf will generalize onto *g* factor. Although both *g* and Gf seem to constitute very close constructs, and the lack of a respective SLODR effect is likely also in the case of *g*, answering this question requires a future study that will include a broader battery of intelligence tests.

Taking all of the above evidence into consideration, the present study brings a new line of evidence for the understanding of fluid intelligence in terms of the effectiveness of working memory processes (including active control and integration of information necessary for constructing adequate mental models) that subsume the correct abstract reasoning in novel situations. The fact that, unlike other potential mechanisms underlying Gf (e.g., processing speed), WMC is immune to the SLODR effect, strongly suggests that WM is more closely linked to Gf than are these other mechanisms.

References

Akaike, H. (1974). A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control, 19*, 716–723.

Burgess, G. C., Gray, J. R., Conway, A. R. A., & Braver, T. S. (2011). Neural mechanisms of interference control underlie the relationship between fluid intelligence and working memory span. *Journal of Experimental Psychology: General, 140*, 674–692.

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices test. *Psychological Review, 97*, 404–431.

Chuderski, A. (2013). When are fluid intelligence and working memory isomorphic and when are they not? *Intelligence, 41*, 244–262.

Chuderski, A. (2014). Relational integration task explains fluid reasoning above and beyond other working memory tasks. *Memory & Cognition*, *42*, 448–463.

Chuderski, A., & Andrelczyk, K. (2015). From neural oscillations to complex cognition: Simulating the effect of the theta-to-gamma cycle length ratio on analogical reasoning. *Cognitive Psychology, 76*, 78–102.

Chuderski, A., Taraday, M., Nęcka, E., & Smoleń, T. (2012). Storage capacity explains fluid intelligence while executive control does not. *Intelligence, 40*, 278–295.

Cleveland, W. S. (1981). LOWESS: A program for smoothing scatterplots by robust locally weighted regression. *The American Statistician, 35*, 54.

Colom, R., Abad, F. J., Quiroga, M. A., Shih, P. C., & Flores-Mendoza, C. (2008). Working memory and intelligence are highly related constructs but why? *Intelligence, 36*, 584–606.

Conway, A. R. A., Getz, S. J., Macnamara, B., & Engel de Abreu, P. M. J. (2011). Working memory and fluid intelligence: A multi-mechanism view. In Sternberg, R. J., & Kaufman, S. B. (Eds.), *The Cambridge Handbook of Intelligence*. Cambridge, UK: Cambridge University Press.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87–114.

Cowan, N., Fristoe, N. M., Elliott, E. M., Brunner, R. P., & Saults, J. S. (2006). Scope of attention, control of attention, and intelligence in children and adults. *Memory & Cognition, 34*, 1754–1768.

Deary, I. J. (2012). Intelligence. In S. T. Fiske, D. L. Schacter, & S. E. Taylor (Eds.), *Annual review of psychology*, Vol. 63. (pp. 453–482).

Der, G., & Deary, I. J. (2003). IQ, reaction time and the differentiation hypothesis. *Intelligence, 31*, 491–503.

Detterman, D. K., & Daniel, M. H. (1989). Correlations of mental tests with each other and with cognitive variables are highest for low IQ groups. *Intelligence, 13*, 349–359.

Evans, M. G. (1999). On the asymmetry of *g*. *Psychological Reports, 85*, 1059–1069.

Goodwin, G. P., & Johnson-Laird, P. N. (2005). Reasoning about relations. *Psychological Review, 112*, 468–493.

Halford, G. S., Baker, R., McCredden, J. E., & Bain, J. D. (2005). How many variables can humans process? *Psychological Science, 16*, 70–76.

Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences, 21*, 803–864.

Hamel, R., & Schmittmann, V. D. (2006). The 20-minute version as a predictor of The Raven Advanced Progressive Matrices Test. *Educational and Psychological Measurement, 66*, 1039–1046.

Hartmann, P. H., & Reuter, M. (2006). Spearman‘s “Law of Diminishing Returns” tested with two methods. *Intelligence, 34*, 47−62.

Hummel, J. E., & Holyoak, K. J. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychological Review, 110*, 220–264.

Jauk, E., Benedek, M., Dunst, B., & Neubauer, A. (2013). The relationship between intelligence and creativity: New support for the threshold hypothesis by means of empirical breakpoint detection. *Intelligence, 41*, 212–221.

Jensen, A. R. (2003). Regularities in Spearman’s law of diminishing returns. *Intelligence, 31,* 95–105.

Kane, M. J., & Engle, M. J. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review, 9*, 637–671.

Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin, 131*, 66–71.

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods, 7*, 19–40.

Martínez, K., Burgaleta, M., Román, F. J., Escorial, S., Shih, P. C., Quiroga, Mª A., & Colom, R. (2011). Can fluid intelligence be reduced to ‹simple› short-term storage? *Intelligence, 39*, 473–480.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence, 37*, 1–10.

Oberauer, K., Schultze, R., Wilhelm, O., & Süß, H.-M. (2005). Working memory and intelligence – their correlation and their relation: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin, 131*, 61–65.

Oberauer, K., Süß, H.-M., Wilhelm, O., & Sander, N. (2007). Individual differences in working memory capacity and reasoning ability. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp. 49–75). Oxford: Oxford University Press.

Oberauer, K., Süß, H.-M., Wilhelm, O., & Wittman, W. W. (2008). Which working memory functions predict intelligence? *Intelligence, 36*, 641–652.

Orzechowski, J., & Chuderski, A. (2007). *Test analogii obrazkowych.* Unpublished manuscript.

Ragni, M., & Knauff, M. (2013). A theory and a computational model of spatial reasoning with preferred mental models. *Psychological Review, 120*, 561–588.

Raven, J. C., Court, J. H., & Raven, J. (1983). *Manual for Raven’s Progressive Matrices and vocabulary scales (Section 4: Advanced Progressive Matrices)*. London: H. K. Lewis.

Reynolds, M. R., & Keith, T. Z. (2007). Spearman’s law of diminishing returns in hierarchical models of intelligence for children and adolescents. *Intelligence, 35*, 267–281.

Reynolds, M. R., Keith, T. Z., & Beretvas, S. N. (2010). Use of factor mixture modeling to capture Spearman’s law of diminishing returns. *Intelligence, 38*, 231–241.

Snow, R. E., Kyllonen, P. C., & Marshalek, B. (1984). The topography of ability and learning correlations. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 2, pp. 47–103). Hillsdale, NJ: Lawrence Erlbaum Associates.

Spearman, C. E. (1927). *The abilities of man.* London: Macmillan.

Unsworth, N., Redick, T. S., Lakey, C. E., & Young, D. L. (2010). Lapses in sustained attention and their relation to executive control and fluid abilities: An individual differences investigation. *Intelligence, 38*, 111–122.

Unsworth, N., Spillers, G. J., Brewer, G. A., & McMillan, B. (2011). Attention control and the antisaccade task. A response time distribution analysis. *Acta Psychologica, 137*, 90–100.

Vigneau, F., Caissie, A. F., & Bors, D. A. (2006). Eye-movement analysis demonstrates strategic influence on intelligence. *Intelligence, 34*, 261–272.

Viskontas, I. V., Holyoak, K. J., & Knowlton, B. J. (2005). Relational integration in older adults. *Thinking and Reasoning, 11*, 390–410.

ACKNOWLEDGMENTS

The study was supported by project 2013/11/B/HS6/01234 granted to A. Chuderski by the National Science Centre of Poland.